Researchers' Perspective on the Publication of Research Data: Semi-structured Interviews from China

Interview: os_030

1	**Interviewer:** So. Again, thank you very much. For this interview. My first question is: How long have you been working in science?
2	**Researcher:** Erm I started to... I think as a PhD student from 2011. So, it's already around eight years.
3	**I:** Nice. And what is your research about?
4	**R:** Erm, So, my background is more like engineering. So, when I was PhD student, so, my research focused on signal processing. Then four years later, so, my research focused on data analytics. So, move from signal processing to data analytics. So currently, it's about educational data mining. A lot of change, yeah.
5	**I:** Very interesting. And what research data do you work with, exactly?
6	**R:** So, for my, for my own research currently, is like two different data set I am working on. So, the first one is erm is an Open Access data set. So, it's just a varied data set. Was published by erm a British university. So, yeah. So, use//using that data set I just try to develop my own algorithm. Just to try to get a better results. Yeah, so that's the first data set. So//
7	**I:** And what kind of data is it? I am sorry to interrupt you. What, what kind of data.
8	**R:** Erm, for that. That data is multi//multi//multimode data. So basically, that university collected ECG, EEG, GSR and vital data. So try to, just to do some emotion detection. Yeah.
9	**I:** Mhm ((affirmative sound)). Nice. Ok. And the other? Data set?
10	**R:** The other data set is... is collected by our own university. So basically, that data set includes for example online and offline learning. And also, inside and outside classroom behaviours.
11	**I:** Mhm ((affirmative sound)).
12	**R:** Data from students. Yeah.
13	**I:** And have you already published your research data?

14	**R:** Oh... that's tricky for the various data. Because some of the data was created by ourselves. But most of the data belongs to the university. So... you know, the way... even for ourselves we sign an agreement. Just to say, ok this data is just for research. So, yeah, we kind of use that data. Erm, I think in the future, maybe it's possible for us to publish that data. But currently is not, is not published yet.
15	**I:** Okay. And what are your biggest concerns, when publishing research data? Would you have any concerns at all?
16	**R:** For my research, for myself? As I research. Yeah, I'd rather to publish the data. Because you know, it's really difficult to find an Open Access data to do your own research. If it can be published, I am sure it will have for//have for researchers from other countries or from other cities, even in China. But it's just a... more discussion the way's the university, with the department, who have this data or who own this data? Yeah, so it's// it's a lot discussions as well. Yeah.
17	**I:** Yeah. And from the legal regulations, would you be allowed to publish?
18	**R:** Erm I would say, it's possible. But erm I//I//I need to//to discuss this issue with different departments. Yeah. So, but that's the plan. So, we hope in the future we can publish our... Maybe even, just our... a small size of this data. A small sample of this data. Yeah.
19	**I:** Mhm ((affirmative sound)). And do you work with personal or sensitive data?
20	**R:** Personal… I//In the future, we are going to collect some personal data ourselves. So, using some variables. But currently, not yet.
21	**I:** Not yet. OK.
22	**R:** Yeah.
23	**I:** And, erm... already said about that. What information would you need, to make the process of publishing easier? What info//Yeah. What information or maybe what tools would you need? So you can make your data public.
24	**R:** Erm... Let's say in the future, if I am going to collect my own data.
25	**I:** Yes.
26	**R:** So... let me think about it. So, I think... So, kind of, erm... How to make they ask our issues to solve. I think that's a first thing just some kind of guideline to tell the researchers, if you are going to publish it, a data set. Erm... Basically, how many different steps you need to do. Is there

anything you need to pay attention? So that's the first thing. And the second thing is, I think currently they ask to publish journal paper or//or workshop paper. So, erm actually publish some paper they//they//they just are try to publish a journal paper or//or academical papers something like that. So just to give a basic idea about, what kind this data can//what kind//what kind of different things this data set can do. And so, just like a baseline, just a offer to other researchers. So, I think that's the second thing. May help. So... third thing, erm... Yeah. I think that those are two most important thing, for me.

| 27 | **I:** Mhm ((affirmative sound)). |

| 28 | **R:** Just a, just... Yeah. |

| 29 | **I:** Perfect. And do you know, erm if research data in you discipline is more or less published in other countries? |

| 30 | **R:** Erm In//in my own research. Hm... No, basically the most data sets are not published. So, I think several years ago there//there was a data set. But it's just like less than one hundred instruments. So that's, I think that's the... Oh, currently there are three, sorry. Three data sets are published. So. But they're small pattern//pattern size, data, the... sample size, yeah. |

| 31 | **I:** Mhm ((affirmative sound)) |

| 32 | **R:** For example, as they are using they//they use like... If they use a whiteboard sensors to collect the data, then the students, the samples is less than one hundred students. So... if a school choose larger sample size, then it's just a use, one single data for number WIFI data, or... Yeah, most//most data sets they are the true data size, just including for number WIFI data. |

| 33 | **I:** Mhm ((affirmative sound)). |

| 34 | **R:** The number come from university campus. |

| 35 | **I:** Mhm ((affirmative sound)). But do you know if in other countries they are publishing data from your, like in your field, in your discipline? |

| 36 | **R:** Erm... I, currently I found three: one is from, from the United States. So, the other two are from China. Yeah. |

| 37 | **I:** Okay. |

| 38 | **R:** That's really the size I found currently. |

| 39 | **I:** Perfect. Oh, that was already my last question ((laugh)). Thank you very much. |